



Gag\_AF110965\_BW\_mod

ATGGGCAGCCGCGCAAGCATCTGCGGGCGCAAGCTGGACGCCCTGGAGCGCATCCGCC  
TGCAGCCCGGGCGCAAGAAGTGCTACATGAGAACCTGGCTGGCCAGCCCGAGCT  
GGAGAAAGTCGCCCTGAACCCGGCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATC  
CGCCAGCTGCACCCGCCCTGAGACCGCAGCGAGGAAGCTGAAGAGCCTGTTCAACACCG  
TGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCCCTGGA  
CAAGATCGAGGAGGAGCAGAACAAAGTGGCAGCAGAACAGATTCAGCAGGGCGAGGCCCGAC  
AAGGGCAAGGTGAGGCCAGAACACTACCCCATCTGCAAGAACCTGAGGGCCAGATGGTGACCC  
AGGCCATCAGCCCCGACCCCTGAAACGCCCTGGTGAAGGTGATCGAGGAGAACGCCCTTCAG  
CCCCGAGGTGATCCCCATGTTACCGCCCTGAGCGAGGGCGCACCCCCCAGGACCTGAAAC  
ACGATGTTGAACACCGTGGGGCCGAGCGGCGCATCGAGATGCTGAAGGACACCATCA  
ACGAGGAGGCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCCGCCCATGCCCGG  
CCAGATGCGCGAGCCCCGCGGAGCAGCATCGCCGGCACACCAGCACCTGAGGAGCAG  
ATCGCCTGGATGACAGCAACCCCCCATCCCGTGGGCGACATCTACAAGCGGTGGATCA  
TCTGGGCAACAGATCTGCGGATCTACAGCCCCGTGAGCATCTGGACATCAAGCA  
GGGCCCCAAGGAGCCCTTCCGCGACTACGTCGGACCGCTTCAAGAACCTGCGCGCGAG  
CAGAGCACCCAGGAGCTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGCCAAC  
CCGACTGCAAGACCATCCTGCGCCTCTGGCCCCGGCGCCAGCTGGAGGAGATGATGAC  
CGCCTGCCAGGGGTGGCGGGCCCAGGCCAAGGGCCCGTGTGGCGAGGCATGAGC  
CAGGCCAACACCAAGGTGATGATGCCAGAACGCAACTTCAAGGGCCCCGGCGCATCGTCA  
AGTGTCTCAACTGCGCAAGGAGGGCACATCGCCCGCAACTGCGCGCCCCCGCAAGAA  
GGGCTGCTGGAAGTGCAGGCAAGGAGGGCCACCGAGATGAAGGACTGCACCGAGCGCCAGGCC  
AACTTCTGGGCAAGATCTGGCCAGCCAACAGGGCCGCCGCAACTTCTGAGAGCC  
GCCCGAGGCCACGGCCCCCGCCAGAGAGCTTCCGCTTCAGGGAGACCACCCCCGGCCA  
GAAGCAGGAGAGCAAGGACCGCGAGACCCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGAC  
CCCCCTGAGCCAGTAA

**FIG. 1**



Gag\_AF110967\_BW\_mod

ATGGGCGCCCGGCCAGCATCTCGCGGGGAGAACGCTGGACAAGTGGAGAAGATCCGCC  
TGCCTCCGGCGCAAGAACGACTACATGCTGAAGCACCTGGTGTGGGAGGCCGAGCT  
GGAGGGCTTCGCCCTGAACCCGGCCTGCTGGAGACGCCAGGGCTGCAAGCAGATCATG  
AAGCAGCTGCAGCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCTGTACAACACCG  
TGGCCACCCCTGACTGCGTGACCGCCGATCGAGGTCCCGACACCAAGGAGGCCCTGG  
CAAGATCGAGGAGGAGCAGAACAAAGTCCCAGCAGAACCCAGCAGGCCAGGAGGCCAC  
GGCAAGGTGAGCCAGAACTACCCATCGTCAGAACCTGCAAGGCCAGATGGTCAACCAGG  
CCATCAGCCCCCGCACCTGAACGCCCTGGTGAAGGTGATCGAGGAGAACGGCTTCAGCCC  
CGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCAAGGACCTGAAACAG  
ATGTTGAACACCGTGGCGCCACAGGCCCATGAGATGCTGAAGGACACCATCAACAG  
AGGAGGCCCGAGTGGGACCCCTGCAACCCGTGCAAGGCCGCCCCGTGCCCCCGGCCA  
GATGCGGACCCCCCGCGCAGCGACATGCCGGCGCACCAGCACCTGCAGGAGCAGATC  
GCCCTGGATGACCAAGCAACCCCCCGTGGCGCACATCTACAAGGGTGGATCATCC  
TGGCTGAAACAGATCGGGATGTCACACCCCGTGAAGCATCTGGACATCCGCCAGGG  
CCCCAAGGAGCCCTTCCGCGACTACGTGAAACGGCTTCTCAAGACCCCTGGCGCCAGCAG  
GCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCAACCCG  
ACTGCAAGACCATCTCGCCGCTCTGGCCCCGGCCACCCCTGGAGGAGATGATGACCGC  
CTGCCAGGGCTGGCGGGCCACAAGGGCCCGTGTGGCCAGGGCATGAGCCAG  
GCCAACAGCGTGAACATCATGATGCAAGAACCTCAAGGGCCCCGGCGAACGTCA  
AGTGTCTCAACTGCGGAAGGAGGCCACATGCCAAGAACCTGCGGCCACCCGCAAGAA  
GGGCTGCTGGAAGTGCAGGAGGCCACAGATGAAGGACTGCAACCGAGCGCAGGCC  
AACTTCTGGCAAGATCTGGCCAGGCCACAAGGGCCGCCGGCAATTCTGCAAGAAC  
GCAGCGAGCCCCGGCCACCGTGGCCACCGCCCCCGCCAGAGCTTCCGCTTCGA  
GGAGACCAACCCCGCCCCCAAGCAGGAGCCAAGGACCGCAGGCCCTACCGCGAGGCCCTG  
ACCGCCCTGCGCAGCCTGTTGGCAGCGGGCCCTGAGGCCAGTAA

**FIG. 2**



**Env\_AF110968\_C\_BW\_opt**

--> **signal peptide (1-81)**

ATGCCGCTGATGGGCATCTGAAGAACTACCAAGCAGTGGTGGATGTGGGGCATCTGGGCTCTGGATGCTGATCA  
 ✓--> gp120(140/160)(82)  
 TCAGCAGCGTGGTGGCAACCTGTGGGTGACCGTGTACTACCGCGTCCCCGTGGAGGAGGCCAACGACCCCT  
 GTTCTGCACCAAGCGACGCCAACGGCTACAGAGACCGAGGTGACAACGTGTGGGCCACCCACGCCCTGCGTGC  
 GACCCCCAACCCCAGGAGATCGTGTGGAGAACGTGACCGAGAACCTCAACATGTGGAAGAACGACATGGTGACC  
 AGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGCGTGA  
 CCTGAAGTGCCTGACCGAACGTGAACGCCAACACATCAACAGCATGATCGACAACAGAACAGAACAGGAGATGA  
 AACTGCAGCTTCACGTGACCCACCGAGCTGCCGACCGAAGCAGGAGGTGACGCCCTGTTCTACGCCCTGAC  
 TGGTGCCCCCTGCAGGGCAACACAGCAACGAGTACCGCCTGATCAACTGCAACACAGGCCATCACCCAGGCC  
 CCCCCAAGGTGAGCTTCAGCCCATCCCCATCCACTACTGACCCCCGCCGTACGCCATCTGAAGTGAACAAAC  
 CAGACCTTCAACGGCACCGGCCCTGCAACAAACGTGAGCAGCGTGCAGTGCGCCACGGCATCAAGCCGTGGTA  
 GCACCCAGCTGCTGTAACCGCAGCCTGGCAAGGGAGATCATCATCCGAGCGAGAACCTGGCAACACAGC  
 CAAGATCATCATCGTGCAGCTGAACAAAGCCGTGAAGATCGTGTGCGTGCCTTCAACACCCCGAACAG  
 GTGCGCATGCCCGGCCAGACCTTCTACGCCACCGCGAGATCATCGGCACATGCCAGGCCACTGCA  
 TCAACAAAGACCGAGTGGAACAGCACCTGCAGGGCGTGAAGCAAGAGCTGGAGGAGCACCTGAGCAAGAGGCC  
 CAAGATTCGAGGCCAGCAGCGGCCGACCTGGAGATCACCAACCCACAGCTTCAACTGCCCGGGGAGTTCTTCTAC  
 TGCGACACCAGCCAGCTTCAACAGCACCTACAGCCCAAGCCTTAACAGGCCAGGAGAACAGCTGAACGGCACCA  
 TCACCATCACCTGCCCATCAACGAGATCATCAACATGTGGCAGAAGGTGGCCGCCATGTACGCC  
 CGCCGGCAACCTGACCTGCGAGGACAACATCACCGCCTGCTGCTGACCCGCCAGCGGCCAGGCC  
 GACACCGAGATCTCCGCCCGGCCGCGACATGCGCAGCAACTGGCCAACAGAGCTGACAGTACAAGGTGG  
 gp120(1512)<--\v-->(1513)gp41  
 TGGAGATCAAGCCCTGGCGTGGCCCGACCGAGGCCAAGCGCCGCGTGGAGGAGCGCAGAGCGGCCGTGG  
 CATCGGCCCGTGTCTGGCTTCTGGGCCGCCGAGCACCATGGGCCGCCAGCATCACCTGACCGTG  
 CAGGCCGCCGTGCTGCGAGCACCATGGGCCGCCAGCACCATGGGCCGCCAGCATCACCTGACCGTG  
 TGCGCAGCTGACCGTGTGGGCATCAAGCAGCTGCGAGGCCGATCCCTGGCGTGGAGCGTACCTGAAGGACCA  
 GCAGCTGCTGGCATCTGGGCTGCGAGGCCAGCTGATCGACCCGCCGTGCCCTGGAACAGCACGACTGGAC  
 AACCGCAGCCACGAGATCTGGACAACATGACCTGGATGCGAGTGGGCCAGGAGATCAACAACTACACCGACA  
 CCATCTACCGCCTGCTGGAGGAGGCCAGAACAGCAGGAGAGAACGAGAGAACGCC  
 gp140(2025)<--\v-->  
 GCAGAACCTGTTGGAACTGGTCAGCATACCAACTGCTGCTGGTACATCAAGATCTTACATCATGATCGTGGCG  
 CTGATCGGCCATCATCTCGCCGTGCTGAGCATCGTAACCGCGTGCGCCAGGCCAGCCTGCC  
 TCCAGACCCCTGACCCCCAACCCCGCGAGGCCAGCCCTGGCCGATCGAGGAGGAGGCCAGGCC  
 CGGCCGCCAGCATCCGCCGTGGTGGAGGCCCTCCCTGGCCCTGGCGACCTGCGCAGCC  
 TACCAACGCCGTGCCGACTCATCTGATGCCGCCCGCGTGTGGAGCTGCTGGGCCAGCGGCC  
 TGAAGTACCTGGCAGCCTGGTGCAGTACTGGGCCCTGGAGCTGAGAACGAGGCC  
 CGCCATGCCGTGGCGAGGCCACCGCATCATCGAGTTCATCCAGCGCAGGCC  
 gp160, gp41(2547)<--\v-->  
 CCCCGCCGATCCGCCAGGGCTCGAGGCCGCCATGCCAACATC

**FIG. 3**



### Env\_AF110975\_C\_BW\_opt

--> **signal peptide (1-72)** ATGCAGCTGGCGCATCTGGCAGCTGGCAGCAGTGGTGATCTGGGATCCTGGGCTTCTGGATCTGCAGC  
**gp120/140/160 (72)** GCCTGGCAACCTGTGGGTGACCGTGTACGACGGCGTCCCCGTGGCGCAGGCCAGCACCCCTGTCGCGC  
 CAGCGACCCAAGGCCTACAGGAAGGAGGTGACAACAGTGTGGGCCACCCACGCCCTGGTGCACCACCCACCCAA  
 CCCCAGGAGATCGAGCTGGACAACGTGACCGAGAACCTCAACATGTGGAAGAACGACATGGTGACCATGAC  
 AGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCCCGGTGAAGCTGACCCCCCTGTGCGTGACCTGAAGT  
 CACCAACTACAGCACCAACTACAGCAACACCATGAACGCCACCGACTACAACAACACACCACCGAGGAGATCAAG  
 AACTGCACCTTCAACATGACCACCGAGCTGGCGACAAGAACGACAGCAGGTGACGCCCTGTTCTACAAGCTGGACA  
 TCGTCCCCCTGAACAGCAACAGCAGCAGTACCGCCTGATCAACTGCAACACCAGGCCATCACCCAGGCCCTGCC  
 CAAGGTGAGCTTCGACCCCCATCCCCATCCACTACTGCGCCCCCGCCGGTACGCCATCTGAAGTGCAAGAACAC  
 ACCAGCAACGGCACGGCCCCCTGCGAGAACGTGAGCACCGTGCAGTGACCCACGGCATCAAGCCCCGTGGTGAGCA  
 CCCCCCTGCTGCTGAACGGCAGCCTGGCGAGGGCGGGAGATCATATCCGCAGCAAGAACCTGAGCAACACG  
 CTACACCATCATCGCACCTGAACGACAGCGTGGAGATCGTGTGACCCGCCAACAAACAACACCCGAAGGGC  
 ATCCGCATCGGCCCCGGCAGACCTTCAACGCCACCGAACATCATGGCGACATCCGCCAGGCCACTGCAACA  
 TCAGCGCCCGCAGTGGACAAGGCCGTGAGCGCTGAGGCCAACCTGCGCGAGCACTCCCCAACAGACCAT  
 CGAGTTCCAGCCCAGCAGCGCCGGGAGCTGGAGATCACCACCCACAGCTTCAACTGCCGCGGGAGTTCTTCTAC  
 TGCAACACCAAGCAAGCTGTTCAACAGCAGCTACAACGCCACAGCTACCGCGGCCAGGAGAACAGCAGCATCA  
 TCACCCCTGCCCTGCCGCATCAAGCAGATCATGACATGTTGAGGCTGGCGGCCATCTACGCCCCCCC  
 CGAGGGCAACATCACCTGCAGCAGCAGCATCACGGCCTGCTGGGCCGACGGGGCCTGGACACATCACC  
 ACCGAGATCTCCGCCCCCAGGGCGGCACATGAAGGACAACGGCAGACTGACAAGTACAAGGTTGG  
**gp120 (1509) <--\--> (1510) gp41**  
 AGATCAAGCCCCCTGGCGTGGCCCCCACCGAGGCCAACGGCCGGTGGAGCGCGAGAACAGCAGGCCAT  
 CGGGCCCCGTGATCTTGGCTTCTGGGCCCGGCCAGCAACATGGGCCCGCAGCATCACCTGACCC  
 GCCCCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACCTGCTGCCGACATCGAGGCCAGCAGCACATGC  
 TGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGGTGCTGCCATCGAGCGTACCTGAAGGACCAGCA  
 GCTGCTGGCATCTGGGCTGAGCGGAAGCTGATCTGACCAACACCGTGCCTGGACAGCAGCTGGAGCAAC  
 AAGACCCAGGGCGAGATGGGAGAACATGACCTGGATGAGCTGGACAAGGAGATCAGCAACTACACCGC  
 TCTACCGCCTGCTGGAGGAGAGCCAGAACCCAGCAGGAGCAGAACGAGAACGACCTGCTGGCC  
**gp140 (2022) <--\-->**  
 CAACCTGTGGAGCTGGTCAACATCAGCAACTGCTGCTGGTACATCAAGATCTTCACTCATGATCGTGGCG  
 ATCGGCCCTGCCGATCATTCGCGTGTGAGCATGTCAGCGTGCAGGCCAGGGTACAGCCCC  
 AGACCCCTGACCCCCAACCCCGCGGCCCTGGACCGCTGGGCCATCGAGGAGGAGGGCGGAGCAGGACCG  
 CGCGAGCATCCGCCCTGGTGCAGGGCTTCTGGGCCCTGGCTGGACGACCTGCCAGCCTGTC  
 CACCGCCTGCCGACCTGATCTGGTGAACGCCCGCGTGGTGGAGCTGCTGGGCCAGCAGCCCC  
 AGCGCGGCCGGAGGCCCTGAAGTACCTGGCAGCCTGGTGCAGTACTGGGCCCTGGAGCTGAAGAACAGGCC  
 CAGCCTGTGGACAGCATGCCATGCCGTGGCGAGGGCACCGCAGCATCGAGGTGATCCAGCGC  
**gp160, gp41 (2565) <--\-->**  
 CGCGCCTCTGCAACATCCCCGCCGCGTGCAGGCCAGGGCTTCGAGGCCGCCCTGCAGTAA

**FIG. 4**

O I P E  
AUG 28 2003  
SCT 19  
RECEIVED  
LIBRARY  
UNIVERSITY OF TORONTO LIBRARIES

Gag\_AF110965\_BW\_opt  
ATGGGGCGCCCGGCCAGC AT CCTGCGCGCGCAAGCTGGACGCCTGGAGGCCATCGCCCTGCGCCCCGG  
CGGCAAGAAGTGTACATGATGAAGCACCTGGTGTGGCCAGCCGAGCTGGAGAAGTTGCCCTGAACC  
CGGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATCCGCAGCTGCACCCGCCCTGCAGACCGGC  
AGCGAGGAGCTGAAGAGCTGTTAACACCGTGGCACCCGTACTGCGTGACGAGAAGATCGAGGT[GCG  
C]  
CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAGAGCCAGCAGAACAGATCCAGCAGGCC  
AGGCCGCCACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTCAGAACCTGCAGGCCAGATGGTGAC  
CAGGCCATCAGCCCCGCACCCGTGAAACGCCCTGGTGAAGGTGATCGAGGAGAACGCCCTCAGCCCCGAGGT  
GATCCCCATGTTACCGCCCTGAGCGAGGGGCCACCCCCCAGGACCTGAACACATGCTGAACACCGTGG  
[G T]  
GGGCCACCAGGCCATGCAGATGCTGAAGGACACCATAACGAGGAGGCCAGTGGGACCGCTG  
CACCCCGTGCACGCCGCCCATGCCCGGCCAGATGCGCGAGCCCCGCCAGCAGCACATGCCGCC  
CACCAAGCACCCTGCAGGAGCAGATGCCCTGGATGACCGAACCCCCCATCCCCGTGGGACATCTACA  
AGCG[G]GGATCATCCTGGCCTGAACAAGATCGTGC[G]ATGTACAGCCCCGTGAGCATTGGACATCAAG  
[G]  
CAGGGCCCAAGGAGCCCTCCGCACTACGTGGACCGCTTCAAGACCCCTGCCGCCAGCAGAGCAC  
CCAGGAGGTGAAGAACTGGATGACCGAACCCCTGCTGGTGCAGAACGCCAACCCCAGTCAAGAACCATCC  
TGCGCG[CCTGGCCCCGGGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGGTGGCGCCCCAGC  
[T C]  
CACAAGGCCCGTGTGCCAGGG[G]ATGAGCCAGGCCAACACCGCGTATGAGCAGAACAGAGCAACTT  
[G]  
CAAGGGCCCCG[G]CATCGT[G]AAGTGTCAACTCGGCCAGGAGGGCACATGCCCGCAACTGCCGCC  
CCCCCGCAAGAAGGGCTGCTGGAAAGTGCAGCAAGGAGGGCACCCAGATGAAGGACTGCACCGAGGCCAG  
GCCAACTTCCCTGGCAAGATCTGGCCAGCCACAAGGGCCGCCAGACTTCCGTGAGGCCAGGGAGAGCAAGG  
GCCCAACGCCGCCAGAGCTTCCGCTTCGAGGAGACCACCCCGGCCAGAACAGCAGGAGAGCAAGG  
ACCGCGAGACCCGTGACCGCCTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAGTAA

**FIG. 5**



Gag\_AF110967\_BW\_opt

ATGGGCGCCCGGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAAGATCCGCCTGCGCCCCGG  
CGGCAAGAACGACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGAGCTGGAGGGCTTCGCCCTGAACC  
CCGGCCTGCTGGAGACCCGCCAGGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCGCCCTGCAGACCGGC  
ACCGAGGAGCTGCGCAGCCTGTACAAACACCGTGGCCACCCCTGTACTGCGTGCACGCCGGCATCGAGGT[GCG  
CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAG[TC]CCAGCAGAACGCCAGCAGGCCA  
AGGAGGCGACGGCAAGGTGAGCCAGAACTACCCCATCGTCAGAACCTGCAGGGCCAGATGGTGCACCAG  
GCCATCAGCCCCCGCACCTGAACGCCCTGGGTGAAGGTGATCGAGGAGAACGGCTTCAGCCCCGAGGTGAT  
CCCCATGTTACCGCCCTGAGCGAGGGGCCACCCCCCAGGACCTGAACAC[CATGG]TGAAACACCGTGGCG  
GCCACCAGGCCCATGCAGATGCTGAAGGACACCATAACGAGGGCCCGAGTGGGACCGCCTGCAC  
CCCGTGCAGGCCGGCCCCGTGGCCCCCGCCAGATGCGCACCCCCCGGGCAGCGACATGCCGGGCCAC  
CAGCACCCCTGCAGGAGCAGATGCCCTGGATGACCAGCAACCCCCCGTGGCGACATCTACAAGC  
G[G]TGGATCATCCTGGCCTGAACAAGATCGTGC[G]ATGTACAGCCCCGTGAGCATCCTGGACATCCGCCAG  
GGCCCCAAGGAGGCCCTCCGCACTACGTGGACCGCTTCTCAAGACCCCTGCGCGCCGAGCAGGCCACCCA  
GGACGTGAAGAACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCAACCCCAGTCAAGAACCATCCTGC  
GCG[G]CTGGCCCCGGGCCACCCCTGGAGGAGATGATGACCGCCTGCCAGGGGTGGCGGCCAC  
AAGGGCCCGGTGCTGGCCGAGGG[C]ATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACT  
CAAGGGCCCCCG[G]CGCAACGT[G]AAGTGCTTCAACTGCGCAAGGAGGGCACATGCCAAGAACGTGCCGG  
CCCCCGCAAGAAGGGCTGCTGGAAGTGCAGGCAAGGAGGGCCACCGATGAAGGACTGCACCGAGCGCCAG  
GCCAACTCTGGCAAGATCTGGCCAGCCACAAGGGCCGCCGGCAACTTCTGCAGAACCGCAGCGA  
GCCGCCGCCACCGTGCACCGGCCACCGGCCGGCAGAGCTCCGCTTCGAGGAGAACCAACCCGCC  
CCAAGCAGGAGGCCAAGGACCGCGAGCCCTACCGCAGGCCCTGACCGCCCTGCGCAGCCTGTTGGCAGC  
GGCCCCCTGAGCCAGTAA

**FIG. 6**